

Mid-Range Computing Working Group Report and Next Steps

Outline

- Milestones
- Completion of Phase I
 - MRC WG Document on initial assessment and findings
 - Recommendations for a path forward
- Phase II
 - Issues
 - Options
 - Next Steps

Milestones

Fall 00	MRC WG is formed
January 01	Phase I Work to define the process
Spring 01	First Plan of activities and schedule Lecture series/Publicity, Identify key users through web-based survey Division Directors buy in, Retreat with key users and technical experts to define architecture Recommendations to upper level LAB management.
May 01	<i>Presentation to CSAC of plan and schedule</i>
Summer 01*	New Plan – Top-down approach and concrete proposal Defining more clearly costs and schedules associated with promoting alvarez to an MRC facility, Executive summary with Bill McCurdy, Workshop with key-users.
August 01	<i>Presentation to CSAC of new plan</i>
September 01	Completion of Phase I MRC Document and Recommendation for a path forward
October 01	Phase II Meeting with Bill McCurdy Issues, Options, Next Steps
January 02	Workshop with key users and experts

** Positive feedback from the Lab Management, and several discussion among MRC-Executive members lead to the “New Plan” and then Phase II*

Phase I: MRC-WG Document

An Institutional Scientific Mid-Range Computing Resource for Berkeley Lab

A report compiled by the Mid-Range Computing Working Group of the Computing and Communications Services Advisory Committee and the Information Technologies and Services Division:

Paul D. Adams, Physical Biosciences

Jon Bashor, Computing Sciences

Ali Belkacem, Chemical Sciences

Alessandra Ciocio, Physics

Kenneth H. Downing, Life Sciences

Gary Jung, Information Technologies and Services

James F. Leighton, Information Technologies and Services

Alexander “Sandy” Merola, Information Technologies and Services

Douglas L. Olson, Nuclear Science

John W. Staples, Accelerator and Fusion Research

Shaheen Tonse, Environmental Energy Technologies

Michel A. Van Hove, Materials Sciences

Tammy S. Welcome, NERSC

Phase I: MRC-WG Document (cont'd)

Executive Summary

“As the role and contributions of high-performance computing continue to increase in significance, Berkeley Lab scientists are seeking out potential advantages provided by more powerful computing resources. These resources range from small clusters developed independently by Lab groups to such high-performance systems as those provided by NERSC.

Based on these indicators, a CSAC-ITSD working group has investigated whether an institutional mid-range computing resource would be appropriate and/or sustainable for Berkeley Lab. This report represents the culmination of the first stage of the group’s work. The working group has identified various options for implementing an institutional mid-range computing resource and identified related financial considerations. The next step is to initiate discussions of such a resource with senior Lab management and the pool of potential users at the Laboratory. Those discussions, together with the information already collected, will then determine the appropriate path forward.”

Phase I: MRC-WG Document (cont'd)

Is an Institutional Mid-Range Computing Resource Appropriate for Berkeley Lab?

The goal of the MRC Working Group is to determine whether an institutional Mid-Range Computing resource is:

- a) Appropriate
- b) Sustainable

“The MRC Working Group made up of CSAC and ITSD members has been assessing whether there is sufficient need and support for such an institutional resource among Berkeley Lab researchers, and to identify additional investments, if any, that Berkeley Lab should make in mid-range computing capabilities.”

The process involves:

- a) Discussion on Possible Options
- b) Find out how mid-range computing has been and is done at LBNL
- c) What, if any, mid-range computing resources are available at other DOE labs
- d) Cost Estimates for different options
- e) Identify potential users and assess user requirements
- f) Define a Viable Financial Model

Phase I: MRC-WG Document (cont'd)

Two Critical Components for Success

Usefulness

To be useful and succeed the mid-range facility:

- a) Should respond to the needs of a broad range of users
- b) Should provide a computing resource that is significantly more powerful than a system that an individual researcher or group could obtain.

It should be readily available, it should have a high turnaround rate, it should have a configuration that responds to the needs of users and it should be relatively easy to use.

- c) Should be perceived by a scientist owning a small cluster as a major step up in terms of advanced computing power and software
- d) Should be upgradeable
- e) Should be much more cost-effective than owning a small cluster
- f) Should be operated in an expert manner
- g) Should be responsive to user needs, requests and input

Commitment

- There should be a clearly expressed need by scientists (and concomitant involvement), a strong commitment from the scientific divisions, and a strong commitment from Lab management.

Phase I: MRC-WG Document (cont'd)

History and Current Status of High-Performance Computing at Berkeley Lab

LBL has shown an interest in MRC:

- Increased usage of NERSC

Since the arrival of NERSC, LBNL has grown from being a minor NERSC user and become the largest Laboratory user of NERSC

- The T3E program at Berkeley Lab demonstrated increased interest

fy'98 was 50K, fy'00 was 191K (total hours allocated)

- MRC can serve as a stepping stone to NERSC
- Growing number of small clusters & SMPs
- PDSF: a Mid-Range Computing Success Story

Cooperative model (Physics, Nuclear Science and NERSC), reliable, well supported, expandable

Phase I: MRC-WG Document (cont'd)

What Are Berkeley Lab's MRC Options

- 1) Providing access to the Lab's newly installed 160-processor cluster named “alvarez,” perhaps with an upgrade
- 2) Contracting for access to computing resources from NERSC, as was done under a special three year program
- 3) Procuring an additional computing resource
- 4) Outsourcing mid-range computing resources
- 5) Making no change at this time

A Financial Model for Institutional Mid-Range Computing

- The financial model must take into account the fiscal realities of Berkeley Lab.
- A viable financial model would involve strong commitment and funding up front from at least several scientific programs and divisions, in conjunction with a contribution from Lab overhead funds.

Phase I: Recommendations for a path forward

Letter to Sandy, September 25, 2001

Summarizing:

- The goal
- The findings
- The two critical issues

Conclusions and recommendations for a path forward:

“Our initial assessment indicates that there is enough interest in the Laboratory for a mid-range computing resource. However, we did not assess the level of commitment that would warrant the viability and usefulness of such a resource.”



As a path forward we recommend first a one-on-one contact with potential mid-range computing users and scientist currently using small cluster machines. The initial discussion should result in the organization of a workshop that will bring together these users to define the need and the level of commitment. The workshop should also involve computer architecture expert to assess system requirements and a viable financial model. If there is sufficient interest, a procurement process and a sustainable financial model would be finalized.

Finally, although with this report, the initial task of this group comes to a completion, most members of the MRC Working Group will gladly assist in organizing the workshop, if the recommendation of the path forward proposed here is endorsed by computing sciences.”

Phase II

Phase II started ...

Meeting with Bill McCurdy

- Summary of MRC Working Group report (Sandy)
- Discussion centered on a realistic assessment of the options available, funding models, sustainability and technical solutions

Phase II: Issues

- There is a growing number of small clusters at LBNL
- Can a MRC facility replace these clusters?
- There must be an added value to central management before divisions/projects would be willing to give up control of owning/running their own systems
 - Fiscal Advantage
 - Cybersecurity
 - Intelligent Scheduling hard but guaranteed access to users who have contributed hardware
 - System administration expertise and purchase of software licenses
 - Physical Environment for the Hardware
 - Fungibility of resources
- ITSD must build expertise to provide added-value
 - Learning process from ground zero.
 - Advantages if NERSC would manage the cluster

Phase II: Issues (cont'd)

- Under any approach, there is an institutional startup cost for MRC, whether it is accomplished through:
 - Alvarez
 - ITSD support of existing systems
 - Small Institutional MRC (whether Cluster or SMP)
- A combined and shared MRC could be managed to provide a more powerful resource than the same capability owned and controlled individually.
- Berkeley Lab management must see an institutional advantage in order to allocate overhead funds

Phase II: Options

- Build on existing MRC
Alvarez or PDSF
- Procure an Institutional MRC
Should it be Cluster or SMP?
- Provide system support
As a gradual mechanism to create a shared resource
- Create an agreement with NERSC
- No offering at this time
Acceptable

Options: Build on existing MRC

PDSF

- PDSF model was a one-time solution funded by DOE/Nuclear Science/Physics
- PDSF is based on program agreement with NERSC
- Politically sensitive to expand for Lab-wide machine
- Not available, in principle, to LBNL divisions whose funds are not primarily DOE

It is difficult for LBNL to build on the PDSF

Alvarez

- This architecture is attractive to MRC WG
- Originally considered evolving into an MRC facility
- That's not where it is presently heading

Dual purpose is difficult

- (grids, not production environment, NERSC investigating technology)

Options: Procure an Institutional MRC

- Interest on the part of the MRC Working Group indicates that acquiring a cluster might have more utility than a small SMP

Costs are Substantial:

16 cpu cluster acquire	\$100K
------------------------	--------

Application License	\$? K
---------------------	--------

First Year and Ongoing	\$? K
------------------------	--------

- Providing a recharge facility is doomed to failure so overhead will be required

Or

- A number of divisions could contribute to the startup costs resulting in:
 - Reduced contribution of overhead by Lab
 - Increased programmatic ownership, a component of long term success
 - Divisions involved must make a long-term commitment to the MRC

Options: System Support as a gradual mean of creating an Institutional Resource

- Individual divisions would be helped in acquiring a MRC. LBNL overhead might fund a system manager for a share programmatic machine
- Consider a number of divisions financing the machine and Lab overhead paying for its operation
- Perhaps, an Institutional Resource might be created if part of the machine(s)' costs were financed by the Lab and made available to other divisions
- Fungible resource could allow building/sharing of a larger machine given the future divisional investments

Options: Create an agreement with NERSC

Economies of scale

- NERSC, already experts in MRC machines, could operate a lab-wide MRC machine
- This model was successful in the past
- A multi-year agreement with NERSC would be established

Could provide 24/7 support

Could hire cluster Linux experts and retain them for long enough periods to be useful

Options: Do nothing

Next Steps

- LBNL needs a Mid-Range Computational Facility
- Good scheduling software is not available but it is necessary
- Guaranteed full access especially if hardware funded by individual users
- Lab-wide cluster will provide better cybersecurity

Options

- Procure an Institutional MRC
Cluster or SMP?
- System support
As a gradual mechanism to create a shared resource
- Create an agreement with NERSC

Next

- Identify PIs who are really interested
Smaller, 20-node clusters are now being purchased by divisions
Must in turn convince their Division Directors to support a facility
Strong push from at least three Div Dir is needed
- Workshop scheduled for approx. January 2002
Must choose the preferred architecture
Find a workable funding model